

Klasifikasi Dana Hibah Usaha Mikro Kecil Dan Menengah dengan Metode Naïve Bayes

Ucta Pradema Sanjaya, Teguh Pribadi, Ifnu Wisma Dwi Prastya

uctapradema@unugiri.ac.id, pribadi.teguh@unugiri.ac.id, ifnuprastya@unugiri.ac.id

Universitas Nahdlatul Ulama Sunan Giri

Informasi Artikel

Diterima : 17 Sep 2022

Direview : 16 Nov 2022

Disetujui : 2 Des 2022

Kata Kunci

Klasifikasi, Naïve Bayes,
Dana Hibah, UMKM

Abstrak

Pandemi covid-19 yang melanda membuat pengusaha mengalami melambatnya perekonomian. Untuk menstimulus perekonomian serta memberikan ketahanan terhadap pengusaha maka pemerintah memberikan dana hibah untuk pengusaha Usaha Micro Kecil dan Menengah. Pemberian dana hibah untuk Usaha Micro Kecil dan Menengah terkadang terdapat masalah dalam pembagiannya. Dikarenakan terdapat permasalahan tersebut maka perlu adanya model data mining dalam menangani masalah tersebut. Data mining bentuk disiplin ilmu yang memiliki 5 peran antar lain metode klasifikasi. Pada metode klasifikasi yang menggunakan peluang yang ciri perhitungannya adalah metode Naïve Bayes. Metode Naïve Bayes sudah banyak digunakan untuk mengklasifikasikan beberapa penelitian terkait dengan ekonomi, kesehatan, dan lain sebagainya. Dari penggunaan Naïve Bayes, maka akan di evaluasi dengan X-Cross validation/ K-Fold validation. Dari perbandingan penggunaan fold validation maka nilai akurasi terbesar terdapat pada nilai 3 fold validation dengan nilai akurasi sebesar 95,96% dan untuk nilai recall pada percobaan metode fold validation semuanya mendapatkan nilai 100%. Nilai presisi paling tinggi pada percobaan 3 fold validation yaitu sebesar 87,96%.

Keywords

*Classification, Naïve Bayes,
Government Assistance,
UMKM*

Abstrak

The COVID-19 pandemic that has hit has made entrepreneurs experience a slowdown in the economy. To stimulate the economy and provide resilience to entrepreneurs, the government provides grants for Micro, Small and Medium Enterprises entrepreneurs. Grants for Micro, Small and Medium Enterprises sometimes have problems in their divisors. Because there are these problems, there is a need for a data mining model in dealing with these problems. Data mining is a form of discipline that has 5 roles, including classification methods. In the classification method that uses opportunities whose characteristic of calculation is the Naïve Bayes method. The Naïve Bayes method has been widely used to classify several studies related to economics, health, and so on. From the use of Naïve Bayes, it will be evaluated with X-Cross validation / K-Fold validation. From the comparison of using fold validation, the largest accuracy value is found at the value of 3 fold validation with an accuracy value of 95.96% and for the recall value in the fold validation method experiment, all of them get a value of 100%. The highest precision value in the 3 fold validation experiment was 87.96%.

A. Pendahuluan

Sektor bisnis di Indonesia masih banyak yang tergantung dengan Usaha Menengah Micro dan Kecil (UMKM) merupakan. UMKM menjadi tulang punggung perekonomian di Indonesia yang sekarang menjadikan fokus dalam pemerintahan ini [1]. Pemerintah telah mengucurkan dana bantuan UMKM untuk mengembangkan usahanya yang masih terpuruk dengan kondisi pasca pandemi Covid-19 [2]. Dengan adanya pandemi covid 19 membuat UMKM merasakan guncangan yang sangat berarti dalam perkembangannya. Dengan guncangan ini membuat perekonomian negara Indonesia menjadi melambat. Pemerintah memberikan dana hibah untuk UMKM guna membantu dalam mengembangkan usahanya ketika pandemi . Akan tetapi dalam pemberian dana hibah ini terkadang tidak tepat sasaran. Maka perlunya pengklasifikasian yang jelas dari data yang di kumpulkan dari kemungkinan yang bisa menjadikan bahan referensi dalam pengambilan keputusan yang tepat sasaran.

Penelitian ini bertujuan untuk memberikan ketepatan dalam pengambilan keputusan berdasarkan data yang sudah ada[3]. Data ini di dapatkan berdasarkan analisis berdasarkan data yang di ambil di lapangan dengan cara interview stakeholder yang berkaitan. Dari data yang didapatkan hendak di proses untuk menentukan aspek mana saja yang akan menjadikan penentu variabel yang akan di jadikan acuan dalam dalam perhitungan klasifikasi[2]. Variabel ini sangat mempengaruhi dalam penentuan klasifikasi untuk menilai seberapa tinggi nilai akurasi serta variabel mana saja yang mempengaruhi dalam menentukan jenis klasifikasi.

Dalam peran data mining di bagi menjadi 5 yaitu klasifikasi, klustering, asosiasi, prediksi, estimasi [4]. Klasifikasi adalah sebuah cara dan proses untuk mengetahui karakter dari kelompok data yang di miliki dari data tersebut [5]–[7]. Adapun beberapa metode klasifikasi yaitu C.45, CART, Naïve Bayes, Support Vector Machine, K-nearest neighbors, logistic regression, dan lain sebagainya. metode metode tersebut memiliki karakteristik dalam melakukannya. Naïve Bayes telah digunakan pada beberapa penelitian terkait dengan bidang kesehatan [6] ekonomi[8], bisnis [9], dan lain sebagainya.

Metode Naïve Bayes [10] merupakan metode yang mempunyai karakteristik dalam menghitung peluang yang terjadi dalam keseluruhan data. Naïve Bayes juga biasa di sebut dengan metode statistic berdasarkan dengan nilai peluang, metode ini melihat dari nilai peluang yang akan di hitung berdasarkan masing masing variabel yang akan masuk dalam kelas tersebut. maka dari itu nilai independensi yang ada dari beberapa variabel akan di hitung akan dapat mempengaruhi dalam menentukan kelas yang akan di tentukan.

Untuk mendapatkan jenis variabel serta pemilihan metode dalam penelitian ini, maka harus berdasarkan study literature yang sudah di lakukan sebelumnya.

Pada tahun 2020 Irwan Ibrahim melakukan penelitian tentang [11] perbandingan Algoritma Naïve Bayes serta C4.5 Untuk pengklasifikasian Bantuan Rumah Sehat. Dari hasil penelitian tersebut perbandingan nilai akurasi dari kedua metode tersebut 95,51% untuk C.45 dan 95,61%, nilai tersebut di dapat dari 3 cross validation pada metode Naïve Bayes . dan 9 cross validation untuk metode C.45. Data yang di olah dalam penelitian tersebut sebanyak 114 data.

Tahun 2020 Rizal Amegia Saputra [12] melakukan penelitian tentang Pemerolehan Informasi Pada Algoritma C4.5 Untuk Klasifikasi Penerimaan Bantuan Pangan Non Tunai (BPNT). Penelitian tersebut menggunakan 130 dataset yang akan di klasifikasikan berdasarkan kriteria KEMENSOS. Untuk akurasi penggunaan metode tersebut menghasilkan nilai sebesar 91,54%.

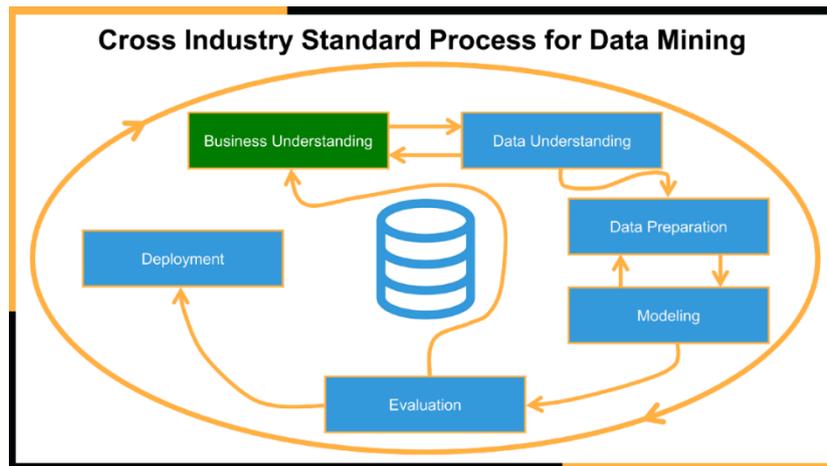
Ahmad Ari Aldino pada tahun 2022 [7] meneliti tentang Analysis of Covid-19 Cash Direct Aid (BLT) Acceptance Using K-Nearest Neighbor Algorithm. Penelitian 711 dataset yang akan di gunakan untuk penelitian menggunakan metode K-Nearest Neighbor. Evaluasi metode menggunakan teknik cross validation. Penelitian tersebut mendapatkan nilai akurasi sebesar 89,68%.

Dari penelitian sebelumnya sudah banyak teknik serta metode klasifikasi yang digunakan yang berkaitan dengan pembagian bantuan dari pemerintah. Banyaknya jenis variabel yang di gunakan dalam penelitian sebelumnya juga bisa di jadikan sebuah saran atau masukan dalam penelitian ini. Selain jenis attribute yang sudah ada, saran atau masukan dari stake holder juga akan di nilai dalam menentukan pemilihan attribute.

Dikarenakan menggunakan propobabilitas atau peluang, sangat relevan untuk untuk menentukan klasifikasi dalam penentuan dana hibah UMKM. Perhitungan Naïve Bayes bisa menjadikan sebuah metode untuk pengklasifikasian. Dan bisa membantu stake holder dalam membantu dalam penentuan dana hibah secara tepat berdasarkan data yang sudah ada dalam pembagiannya.

B. Metode Penelitian

Penelitian ini bisa menggunakan pendekatan CRISP-DM. Gambar 1 menunjukkan 6 (enam) tahap dalam CRIPS-DM. Enam tahapan CRISP-DM antara lain memahami bidang usaha, memahami suatu data, mempersiapkan data, pemodelan, pertimbangan, serta penyebaran [11]. Tahap dalam memahami bidang usaha berisi penentuan target bidang usaha, penilaian keadaan waktu ini, serta penetapan target untuk penambangan data. Pemahaman data, kegiatan penyusunan, evaluasi syarat data, serta dalam pendataan. Persiapan data, informasi yang telah dikumpulkan maka akan diidentifikasi, data perlu dipilih, bisa juga dibersihkan, dan kemudian data dibangun ke dalam bentuk/format yang diinginkan. Persiapan data, juga dikenal sebagai preprocessing data. Pemodelan adalah bentuk diagram alur atau penerapan algoritma untuk menemukan, mengidentifikasi, dan menampilkan pola. Evaluasi, dipakai dalam mengukur pertimbangan suatu model. Kita dapat memperkirakan bentuk yang terbaik dipakai dalam teknik penambangan data. Tahap penyebaran digunakan untuk mengotomatisasi bentuk ataupun dalam mengembangkan suatu aplikasi, terpadu dengan administrasi yang ada ataupun sistem informasi yang layak



Gambar 1 Alur CRIPS-DM

1. Business Understanding

Tujuan dari penelitian ini, mengali informasi yang ada dari sebuah data untuk mencari pengetahuan yang tersimpan pada data tersebut dengan metode Naïve Bayes untuk diketahui nilai dari akuisi pada penelitian tersebut. Untuk mendapatkan model klasifikasi yang akan membantu dinas terkait dalam menentukan pemberian dana hibah untuk UMKM. Di karenakan banyaknya protes warga tentang kurang tepatnya sasaran dalam pemberian dana hibah tersebut.

2. Data Understanding

Penelitian ini menggunakan dataset tersebut dengan pemilihan variabel yang akan di berikan serta di bahas bersama dengan stakeholder yang berkaitan, untuk memilih mana yang terbaik dalam pemberian dana hibah UMKM. Pada penentuan variabel ini perlu pendalaman antara penelitian terkait dengan stakeholder. Dalam penelitian sebelumnya untuk pemberian dana bantuan yang bersumber dari pemerintah yang menggunakan variabel data pribadi seperti jumlah tanggungan, pendapatan, jenis rumah dan lain sebagainya. Akan tetapi itu di rasa bukan terkait dengan UMKM. Di karenakan dana pemberian ini adalah dana untuk pengembangan usaha makana stakeholder dan peneliti memilih variabel data yang terkait dengan keberlangsungan usaha tersebut. pemilihan data ini untuk mengetahui perkembangan usaha yang sudah di laksanakan. Pemberian bantuan ini bertujuan untuk menstimulus usaha agar bisa menjadi lebih kuat atau bertahan dalam pandemi covid-19.

3. Data preparation

Dataset penelitian adalah data yang digunakan untuk uji coba pada penelitian yang dilakukan. Data ini berisi detail penghasilan dari UMKM dan juga total asset yang dimiliki oleh UMKM tersebut. Dataset ini berasal dari Dinas Perdagangan Kabupaten Bojonegoro yang diambil ketika melakukan observasi lapangan. Data tersebut berisi detail data penghasilan harian UMKM, penghasilan mingguan UMKM, data penghasilan bulanan UMKM dan total asset yang dimiliki oleh UMKM. Dari beberapa data diatas akan dijadikan sebagai atribut penelitian yang nantinya akan dimodelkan menggunakan metode Naïve Bayes. Berikut merupakan data yang

didapatkan dalam penelitian dalam bentuk rupiah. Adapun keterangan dari dataset yang digunakan sebagai berikut:

Tabel 1 Dataset yang Sudah Di Pilih

No	Penghasilan	Penghasilan	Penghasilan	Total Aset	Kesimpulan
	Harian	Mingguan	Bulanan		
1.	50000	350000	1500000	4200000	Dapat
2.	75000	525000	2250000	4200000	Dapat
3.	75000	525000	2250000	4200000	Dapat
4.	150000	1050000	4500000	4200000	Tidak dapat
5.	100000	700000	3000000	4200000	Tidak dapat
6.	50000	350000	1500000	4200000	Dapat
7.	50000	350000	1500000	4200000	Dapat
8.	85000	574000	2460000	4200000	Dapat
....	
100	50000	350000	1500000	4200000	Dapat

Data tersebut dipilih setelah berdasarkan penelitian terkait yang biasa pilih untuk mengklasifikasikan. Bentuk bentuk penelitian terkait dengan pemberian bantuan pemerintah sudah banyak dilakukan. Akan tetapi hasil dari pemilihan atribut untuk penelitian ini tidak berdasarkan data pribadi pemilik sebuah usaha UMKM, dipiilih berdasarkan cashflow dan aset yang di miliki yang ada pada usaha tersebut. Dengan tujuan dalam penelitian ini bertujuan supaya pemberian dana hibah terhadap UMKM tepat sasaran, untuk memberikan stimulus pada pemberdayaan UMKM yang terdampak pandemi Covid-19.

4. Modeling

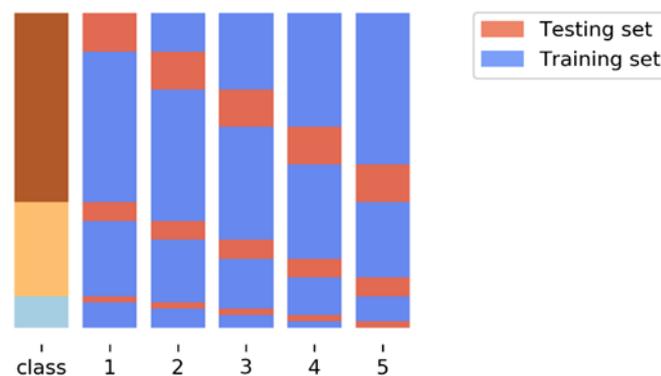
Metode Naïve Bayes mempunyai karakteristik dalam menghitung peluang yang terjadi dalam keseluruhan data. Naïve Bayes juga biasa di sebut dengan metode statistic berdasarkan dengan nilai peluang, metode ini melihat dari nilai peluang yang akan di hitung berdasarkan masing masing variabel yang akan masuk dalam kelas tersebut. maka dari itu nilai independensi yang ada dari beberapa variabel akan di hitung akan dapat mempengaruhi dalam menentukan kelas yang akan di tentukan. Rumusan Masalah pada metode ini ialah [12]:

$$P(H|X) = \frac{P(X|H) P(H)}{P(X)}$$

Dimana X merupakan data nilai class yang belum di ketahui. H merupakan hipotesis dari sebuah data X yaitu class yan spesifik. Untuk nilai probabilitas hipotesis H berdasarkan dengan kondisi sebuah X yang dapat di tulis dengan P(H|X). untuk P(H) sendiri, ialah probabilitas hipotesis dari H. P(X|H) sendiri adalah probabilitas nilai X berdasarkan nilai dari H. P(X) adalah probabilitas X.

5. Evaluasi

Dalam proses pengolahan data, sebelumnya dilakukan proses pembagian data. Proses pembagian data yang dimaksud adalah proses dimana pembagian test data serta training data. Dalam meneliti pembagian data memakai metode *K-Fold Cross Validation*. Metode ini digunakan untuk pembagian test data serta training data yang memiliki tujuan meminimalisir kebetulan pembagian data yang baik seperti ketika menggunakan metode split data[7]. Metode K-Fold Cross validation membagi jumlah dataset sesuai dengan nilai k yang ditentukan [13]. Nilai k digunakan ditentukan oleh peneliti untuk memastikan setiap kelas uji memiliki nilai akurasi. Setelah setiap kelas terhitung akurasinya dilakukan penjumlahan semua akurasi dan dibagi jumlah k yang sudah ditentukan. Adapun skenario yang dilakukan dalam penerapan K-Fold Cross Validation meliputi sebagai berikut



Gambar 2 Validasi Silang

Untuk menentukan bagaimana tentang akurasi atau ketepatan dalam penggunaan metode Naïve Bayes maka menggunakan akan di lakukan teknik validasi silang untuk mendapatkan confusion matrix yang berfungsi untuk mendapatkan nilai recall presisi dan akurasi terhadap metode Naïve Bayes. Perhitungan confusion matrix ini mempunyai rumus di bawah ini [14], [15];

$$akurasi = \frac{TP + TN}{TP + FP + FN + TN}$$

$$Recall = \frac{TP}{FN + TP}$$

$$Presisi = \frac{TP}{FP + TP}$$

C. Hasil dan Pembahasan

Dalam metode Naïve Bayes diusulkan terhadap data yang akan di hitung dengan cara sebagai berikut:

1. Sebagai langkah metode Naïve Bayes untuk mengetahui penghasilan harian sebesar Rp 115.000 penghasilan mingguan sebesar Rp 805.000, penghasilan bulanan sebesar Rp 4.200.000, dan asset yang di miliki sebesar Rp 4.200.000. apakah mendapatkan dana hibah UMKM?

2. Pertama yang dihitung yaitu nilai jumlah kelas yang ada dari keseluruhan data atau dalam rumus di sebut $P(H)$. Di karenakan data yang ada total 100 data maka class yang mendapatkan dana hibah bertotal 74 dan yang tidak mendapatkan dana hibah total 26.
3. Setelah itu mencari perhitungan $P(CI)$, nilai $P(CI)$ ini di dapatkan dari perhitungan seluruh data dibagi dengan nilai total class tersebut. Untuk probabilitas class/ label dapat mempunyai nilai $P(CI)$ 74 di bagi 100 data maka nilainya sebesar 0,74 dan tidak dapat bernilai 0,26. Yang di dapatkan dari 26 di bagi 100
4. Setela mendapatkan nilai $P(X|CI)$, maka di lanjutkan mencari nilai probabilitas tiap variabel terhadap class yang ada seperti dibawah ini:

Tabel 2 Perhitungan $P(X|CI)$ Class Dapat

CI=dapat	
X(harian = 115000)	0.0000
X(mingguan = 805000)	0.0000
X(bulanan = 345000)	0.0000
X(asset=4200000)	0.0000

Tabel 3 Perhitungan $P(X|CI)$ Class Tidak Dapat

CI= tidak dapat	
X(harian = 115000)	0.0385
X(mingguan = 805000)	0.0385
X(bulanan = 345000)	0.0385
X(asset=4200000)	1.0000

5. Setelah mendapatakan nilai tersebut maka akan menghitung nilai $P(X_{total}|dapat)$ dan $P(X_{total}|Tidakdapat)$. Untuk nilai dari $P(X_{total}|dapat)$ bernilai 0,0000 dan $P(X_{total}|tidakdapat)$ bernilai 0.00006. nilai ini didapatkan dari perkalian dari keseluruhan jumlah variabel yang ada.
6. Setelah mendapatkan nilai $P(X_{total}|dapat)$ dan $P(X_{total}|Tidakdapat)$ maka di cari lagi nilai untuk $P(X|Ci)*P(Ci)$ kelas dapat yang bernilai 0,0000 dan $P(X|Ci)*P(Ci)$ kelas tidak dapat bernilai 0,0014.
7. Untuk tahapan terakhir maka akan di bandingkan antara nilai kelas dapat dan kelas tidak dapat. Dari percobaan variabel pada data penghasilan harian sebesar Rp 115.000 penghasilan mingguan sebesar Rp 805.000 , penghasilan bulanan sebesar Rp 4.200.000, dan asset yang di miliki sebesar Rp 4.200.000. maka hasilnya adalah tidak dapat bantuan dana hibah UMKM.

Dalam valiadasi silang peneliti menerapkan *3 fold validation*, *5 fold validation*, *10 fold validation*.

1. Untuk *3 fold validation* menghasilkan confusion matrik

Tabel 4 Hasil Confusion Matrik Dari 3 Fold Validation

	True Dapat	True Tidak Dapat
Pred. Dapat	70	0
Pred. Tidak dapat	4	26

Maka nilai akurasi

$$akurasi = \frac{70 + 26}{70 + 0 + 4 + 26} = 95,96\%$$

Nilai Recall

$$Recall = \frac{26}{26 + 0} = 100\%$$

Nilai presisi

$$Presisi = \frac{26}{4 + 26} = 87,97\%$$

2. Untuk 5 *fold validation* menghasilkan confusion matrix

Tabel 5 Hasil Confusion Matrix dengan 5 Fold Validation

	True Dapat	True Tidak Dapat
Pred. Dapat	66	0
Pred. Tidak dapat	8	26

Maka nilai akurasi

$$akurasi = \frac{66 + 26}{66 + 0 + 8 + 26} = 92\%$$

Nilai Recall

$$Recall = \frac{26}{26 + 0} = 100\%$$

Nilai presisi

$$Presisi = \frac{26}{8 + 26} = 79,29\%$$

3. Untuk 10 *fold validation* menghasilkan confusion matrik

Tabel 6 Hasil Confusion Matrix dengan 5 Fold Validation

	True Dapat	True Tidak Dapat
Pred. Dapat	67	0

Pred. Tidak dapat	7	26
-------------------	---	----

Maka nilai akurasi

$$akurasi = \frac{67 + 26}{66 + 0 + 7 + 26} = 93\%$$

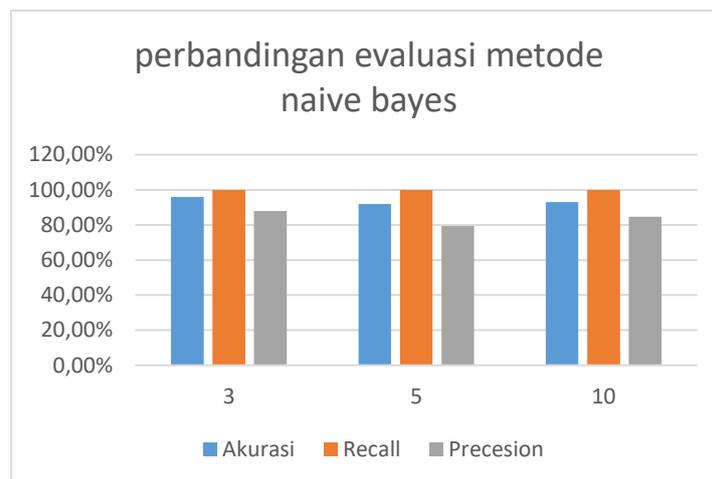
Nilai Recall

$$Recall = \frac{26}{26 + 0} = 100\%$$

Nilai presisi

$$Presisi = \frac{26}{7 + 26} = 84,67\%$$

Dari perhitungan confusion matrik di atas untuk memperjelasnya kami tampilan dalam bentuk gambar grafik di bawah ini:



Gambar 3 Grafik Perbandingan Evaluasi Metode Naive Bayes

D. Simpulan

Berdasarkan penelitian yang sudah di lakukan maka bisa disimpulkan yaitu pada evaluasi menggunakan 3 Fold validation mendapatkan nilai akurasi sebesar 95,96% recall 100% dan presisi 87,96%. 5 Fold validation mendapatkan nilai akurasi sebesar 92,96% recall 100% dan presisi 79,29%. 10 Fold validation mendapatkan nilai akurasi sebesar 93,96% recall 100% dan presisi 84,67%. Dari hasil perbandingan penggunaan fold validation maka nilai akurasi terbesar terdapat pada nilai 3 fold validation dengan nilai akurasi sebesar 95,96% dan untuk nilai recall pada percobaan metode fold validation semuanya mendapatkan nilai 100%. Nilai presisi paling tinggi pada percobaan 3 fold validation yaitu sebesar 87,96%.

E. Ucapan Terima Kasih

Kami berterima kasih kepada Lembaga Penelitian dan Pengabdian Masyarakat Universitas Nahdlatul Ulama Sunan Giri, Dinas Perdagangan Kabupaten Bojonegoro. Yang sudah membantu dalam melaksanakan penelitian ini.

F. Referensi

- [1] Suhery, T. Putra, and Jasmalinda, "Analisis Dampak Covid-19 Terhadap Umkm (Studi Kasus : Home Industri Klepon Di Kota Baru Driyorejo)," *J. Inov. Penelit.*, vol. 1, no. 3, pp. 1–4, 2020.
- [2] N. M. A. Anggraeni Charismanur Wilfarda, Wulan Puspita Ningtiyas, "Kebijakan Pemerintah Dalam Pemberdayaan UMKM Di Masa Pandemi," *J. Gov. Polit.*, vol. 3, no. 1, pp. 47–65, 2021.
- [3] A. Majid, D. Dwi Kurniawan,) Kharisma, and N. Sigit, "Pengaruh Bantuan Presiden Blt Umkm Terhadap Produktivitas Usaha Mikro Kecil Dan Menengah Di Kabupaten Batang," *J. Ilmu Manaj. dan Akunt. Terap.*, vol. 12, no. 3, pp. 333–341, 2021.
- [4] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*. 2006.
- [5] M. M. Alfitri, E. Mashamy, I. Komputer, P. Studi, S. Informasi, and U. D. Ali, "Klasifikasi Data Penduduk Untuk Menerima Bantuan Pangan Non Tunai Menggunakan Algoritma Naïve Bayes," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 4, pp. 1035–1043, 2022.
- [6] M. Alwateer, A. M. Almars, K. N. Areed, M. A. Elhosseini, A. Y. Haikal, and M. Badawy, "Ambient healthcare approach with hybrid whale optimization algorithm and Naïve Bayes classifier," *Sensors*, vol. 21, no. 13, pp. 1–21, 2021.
- [7] A. A. Aldino, R. R. Suryono, and R. Ambarwati, "Analysis of Covid-19 Cash Direct Aid (BLT) Acceptance Using K-Nearest Neighbor Algorithm," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 16, no. 2, p. 193, 2022.
- [8] Pristiyono, M. Ritonga, M. A. Al Ihsan, A. Anjar, and F. H. Rambe, "Sentiment analysis of COVID-19 vaccine in Indonesia using Naïve Bayes Algorithm," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1088, no. 1, p. 012045, 2021.
- [9] L. C. C. Fung, *Advances in Applications of Data-Driven Computing*, vol. 1319. 2021.
- [10] Y. Yuliana, P. Paradise, and K. Kusriani, "Sistem Pakar Diagnosa Penyakit Ispa Menggunakan Metode Naive Bayes Classifier Berbasis Web," *CSRID (Computer Sci. Res. Its Dev. Journal)*, vol. 10, no. 3, p. 127, 2021.
- [11] M. North, *Data Mining for the Masses Second Edition*, vol. 6, no. 2. A Global Text Project Book This, 2012.
- [12] R. Yesputra and A. Sapta, "Prediksi Kelancaran Pembayaran Angsuran Pada Koperasi Dengan Metode Naive Bayes Classifier," *Indones. J. Comput. Sci.*, vol. 6, no. 1, p. 62, 2022.
- [13] C. C. Aggarwal, *Data Mining: The Textbook*. 2015.
- [14] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *J. Ris. Komput.*, vol. 8, no. 6, pp. 219–225, 2021.
- [15] I. Mubarog, A. Setyanto, and H. Sismoro, "Sistem Klasifikasi Pada Penyakit Breast Cancer Dengan Menggunakan Metode Naïve Bayes," *Creat. Inf. Technol. J.*, vol. 6, no. 2, p. 109, 2021.